nature microbiology

ARTICLES

PUBLISHED: 14 NOVEMBER 2016 | ARTICLE NUMBER: 16206 | DOI: 10.1038/NMICROBIOL.2016.206

Single-cell RNA-seq ties macrophage polarization to growth rate of intracellular *Salmonella*

Antoine-Emmanuel Saliba[1,2], Lei Li[2], Alexander J. Westermann[1], Silke Appenzeller[2], Daphne A. C. Stapels[3], Leon N. Schulte[1], Sophie Helaine[3] and Jörg Vogel[1]*

Intracellular bacterial pathogens can exhibit large heterogeneity in growth rate inside host cells, with major consequences for the infection outcome. If and how the host responds to this heterogeneity remains poorly understood. Here, we combined a fluorescent reporter of bacterial cell division with single-cell RNA-sequencing analysis to study the macrophage response to different intracellular states of the model pathogen *Salmonella enterica* serovar Typhimurium. The transcriptomes of individual infected macrophages revealed a spectrum of functional host response states to growing and non-growing bacteria. Intriguingly, macrophages harbouring non-growing *Salmonella* display hallmarks of the proinflammatory M1 polarization state and differ little from bystander cells, suggesting that non-growing bacteria evade recognition by intracellular immune receptors. By contrast, macrophages containing growing bacteria have turned into an anti-inflammatory, M2-like state, as if fast-growing intracellular *Salmonella* overcome host defence by reprogramming macrophage polarization. Additionally, our clustering approach reveals intermediate host functional states between these extremes. Altogether, our data suggest that gene expression variability in infected host cells shapes different cellular environments, some of which may favour a growth arrest of *Salmonella* facilitating immune evasion and the establishment of a long-term niche, while others allow *Salmonella* to escape intracellular antimicrobial activity and proliferate.

Bacterial infections constitute some of the most complex inter-organismic interactions, and it is increasingly recognized that both the pathogen and the invaded host cell can exhibit a large phenotypic heterogeneity during the infection process[1,2]. Examples of phenotypic diversity include variations in bacterial lipopolysaccharide (LPS) composition or intracellular growth status and in host defence mechanisms. This heterogeneity can be a major determinant of disease outcome and the success of anti-microbial treatment[1,3]. *In vivo* studies have shown that bacterial pathogens can exist in different subpopulations, from highly proliferative to non-replicative individuals, a subset of which can 'persist' up to years in the host[4–6]. Heterogeneity in pathogen behaviour is also observed in *ex vivo* infection of monolayers of seemingly homogeneous cultured mammalian cells[7]. Whether and how individual host cells respond to the diverse states of intracellular pathogens is largely unknown.

Global transcriptome studies have been crucial in identifying the signalling pathways whereby mammalian host cells respond to intracellular bacteria[8]. However, standard approaches average gene expression levels from millions of cells, and therefore disregard phenotypic diversity between individual cells. Single-cell transcriptomics using next-generation transcript sequencing (RNA-seq) is now emerging as a powerful tool to profile cell-to-cell variability on a genome-wide scale[9]. Single-cell RNA-seq analyses of LPS-stimulated immune cells revealed an unexpected degree of bimodal expression of seemingly highly expressed immune genes on the population average[10]. This raises the possibility that a range of functional states of immune cells provide different niches for distinct subpopulations of an intracellular bacterial pathogen.

Infections by serovars of *Salmonella enterica* (henceforth *Salmonella*) represent a high global burden, with more than one hundred million cases per year[11]. *Salmonella* is also a primary model organism to study the lifestyle of intracellular bacteria[11]. Macrophages (MΦs) are considered a favourite niche for *Salmonella* to divide and persist within the host organism[4,12]. Recently, a pioneering single-cell RNA-seq study revealed considerable variability in the MΦ response to infection with a single *Salmonella* bacterium[3]. Specifically, differential activation of a key virulence regulator of *Salmonella*, the PhoP/Q two-component system, was linked to a specific gene expression program in the host. However, how MΦs respond to intracellular bacterial growth heterogeneity has remained unexplored, for a lack of experimental tools to combine replication measurements with single-cell transcriptomics[13]. Here, we have used fluorescent *Salmonella* strains reporting bacterial proliferation inside single host cells[5–7] to address how MΦs respond to extreme variations of intracellular pathogen growth. Combining cell sorting and single-cell RNA-seq, we studied global mRNA profiles of individual MΦs harbouring either growing or non-growing bacteria, in addition to those of uninfected bystander cells. Our study suggests that gene expression heterogeneity among infected MΦs creates diverse environments for *Salmonella* to either quietly persist below the radar of intracellular immune surveillance mechanisms or exploit its host for maximal proliferation.

Results

Sorting and single-cell RNA-seq of MΦs infected with *Salmonella*. We infected primary mouse bone marrow-derived MΦs with a *Salmonella* strain carrying an established fluorescence dilution (FD) plasmid that reports bacterial growth after internalization[6,7]. This plasmid provides two different fluorescent markers: constitutively expressed red fluorescent protein (mCherry), which functions as a general marker for the presence of bacterial cells, and conditionally expressed green fluorescent

[1]RNA Biology Group, Institute for Molecular Infection Biology, University of Würzburg, Josef-Schneider-Straße 2, D-97080 Würzburg, Germany. [2]Core Unit SysMed, University of Würzburg, Josef-Schneider-Straße 2, D-97080 Würzburg, Germany. [3]Section of Microbiology, Medical Research Council (MRC) Centre for Molecular Bacteriology and Infection, Imperial College London, Armstrong Road, London SW7 2AZ, UK. *e-mail: joerg.vogel@uni-wuerzburg.de

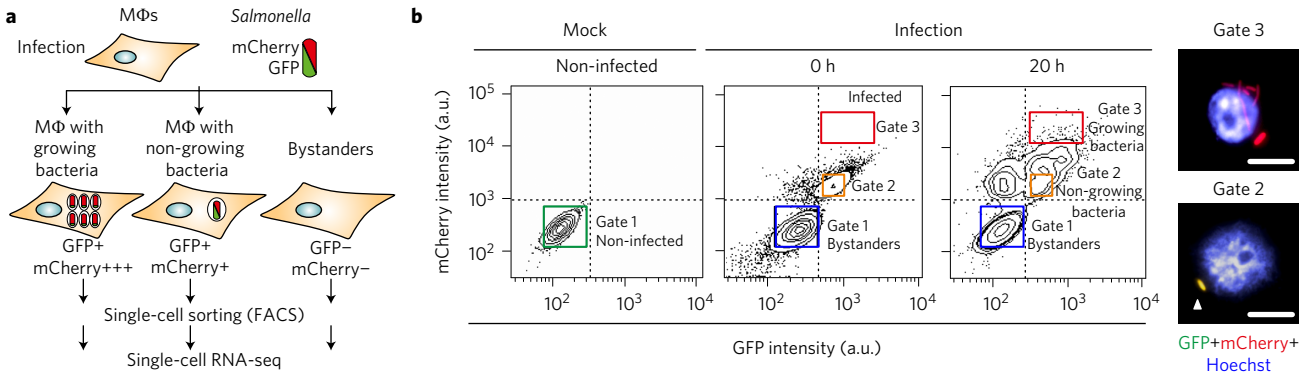

Figure 1 | Experimental strategy to sort single MΦs with different bacterial content. a, Schematic representation of the workflow. MΦs are infected with the intracellular pathogen *Salmonella* Typhimurium strain SL1344 harbouring a dual-colour reporter that comprises a constitutively (mCherry, red) and an arabinose-inducible (GFP, green) fluorescent protein. At 20 h after uptake, a heterogeneous population of host cells is detectable that consists of MΦs with growing bacteria (left), MΦs with non-growing bacteria (middle) and uninfected bystanders (right). FACS coupled to single-cell RNA-seq is used to sort and analyse the different subpopulations. **b**, Representative FC scatter plots of non-infected (mock) and challenged MΦs. Gate 1 captures naive MΦs and bystanders, gate 2 infected MΦs with a single bacterium, and gate 3 MΦs infected with bacteria that have proliferated. Differential bacterial contents were confirmed using fluorescence microscopy, and two representative images are shown (right: Hoechst, blue; GFP, green; mCherry, red). The white arrowhead indicates a non-growing *Salmonella* cell (yellow). Scale bars, 10 μm.

protein (GFP), the dilution of which is used as a proxy for cell division. GFP is expressed from an arabinose-inducible promoter, and its synthesis is terminated by withdrawal of the inducer at the onset of infection. Consequently, while infected MΦs display increasing mCherry signal intensity as bacteria proliferate, they emit a constant GFP signal throughout infection (Fig. 1a). In contrast, MΦs containing non-growing *Salmonella* display constant mCherry and GFP intensities over time, and MΦs containing dead bacterial cells and debris are mCherry-positive but exhibit virtually no GFP signal (data not shown).

We verified, for our experimental set-up, that all bacteria expressed mCherry (Supplementary Fig. 1a,b) and considered non-fluorescent MΦs as bystanders, that is, uninfected host cells that were nonetheless exposed to extracellular bacterial triggers such as the TLR4 ligand LPS. Absence of *Salmonella* from these bystander cells was supported by a lack of a bacterial ribosomal RNA signal in real-time quantitative reverse transcription PCR (qRT-PCR) experiments (Supplementary Fig. 1c). Non-infected naive MΦs from a mock-treated culture constituted the reference control.

We used fluorescence-activated cell sorting (FACS) to isolate single infected and non-infected cells 20 h after *Salmonella* uptake (Fig. 1b), focusing on the extremes of bacterial growth and neglecting intermediate stages. Gentamicin was used to prevent reinfection of MΦs (see Methods). At this time, the immune programs linked to recognition of pathogen-associated molecular patterns (PAMPs) are fully established[10]. We confirmed (Supplementary Fig. 2) that intracellular bacterial growth only starts 6 h post-internalization (h.p.i.), when SPI-2 expression is fully established, and continues at a rate of approximately one division every 3 h (ref. 7). We also confirmed that the gating strategy enabled the selective isolation of MΦs with distinct intracellular bacterial loads by re-analysing sorted cells by fluorescence microscopy (Fig. 1b).

Sorted MΦs were subjected to RNA-seq analysis with the previously described Smart-seq2 protocol[14], which is based on poly(A) capture, template switching and PCR amplification. External RNA Controls Consortium (ERCC) spike-in was added to assess technical noise (see Methods). In total, we processed 64 single cells (4×16 single cells from each population; Fig. 1b and Supplementary Table 1) and filtered out cells with incomplete amplification, insufficient read number or spike-in-dominated libraries (see Methods). In this way, sequencing data from a total of 60 cells remained for further analysis (Supplementary Table 1), representing 15 naive MΦs, 15 bystanders, 16 MΦs with non-growing bacteria and 14 MΦs with growing bacteria. For these cells, >99% of the reads passed our quality-control filters (Supplementary Table 1), and the majority of them (>60%) mapped to protein-coding genes (Supplementary Fig. 3a). On average, $4{,}912 \pm 373$ transcripts were detected in each cell (see Methods) (Supplementary Table 1). The dynamic range of the assay spanned over five orders of magnitude, and single-molecule sensitivity was achieved with efficient detection (>90%) for transcripts with an absolute copy number >10 (Supplementary Fig. 3b,c).

Population structure of infected MΦs based on gene expression patterns. We partitioned the sorted MΦs into different populations based on their transcriptomic similarity, taking an unsupervised approach that is blind to the identity of the originating populations. First, we used the Pearson correlation coefficients associated with hierarchical clustering to measure similarity among the expression profiles of all cells (Fig. 2a). Two major populations corresponding to naive and challenged cells were identified, supporting previous findings that *Salmonella* infection causes a dramatic reorganization of the host transcriptome[8].

To further dissect the two main groups, we applied principal component analysis (PCA) (Fig. 2b) to all 60 cells, including all genes whose expression exceeded technical noise (Supplementary Fig. 3d) (see Methods). This approach identified three subpopulations referred to as groups I, II and III. Group I represents the naive MΦs, and groups II and III represent two subpopulations of challenged MΦs (Supplementary Table 2a). The 15 naive MΦs (group I) clearly segregated from the challenged MΦs along the first PCA component (Fig. 2b), whereas the challenged cells were discriminated along the second PCA component. Group II (28 cells, the second dimension (Dim 2) > 0 (Fig. 2b)) contained all 15 bystander cells, 11 of the 16 MΦs with non-growing bacteria, and two MΦs with growing bacteria. Group III (17 cells, Dim 2 < 0 (Fig. 2b)) was dominated by MΦs infected with growing bacteria (12/14), but also contained MΦs with non-growing bacteria (5/16). Complementary to PCA, *t*-distributed stochastic neighbour embedding (*t*-SNE) was used to search for a more complex cell population organization. However, when projected onto two dimensions, the cells fell into three distinct clusters, following the same pattern obtained for PCA (Supplementary Fig. 4a): naive MΦs segregated from the challenged MΦs along the first dimension and along the second dimension, bystanders and MΦs with non-growing bacteria grouped together

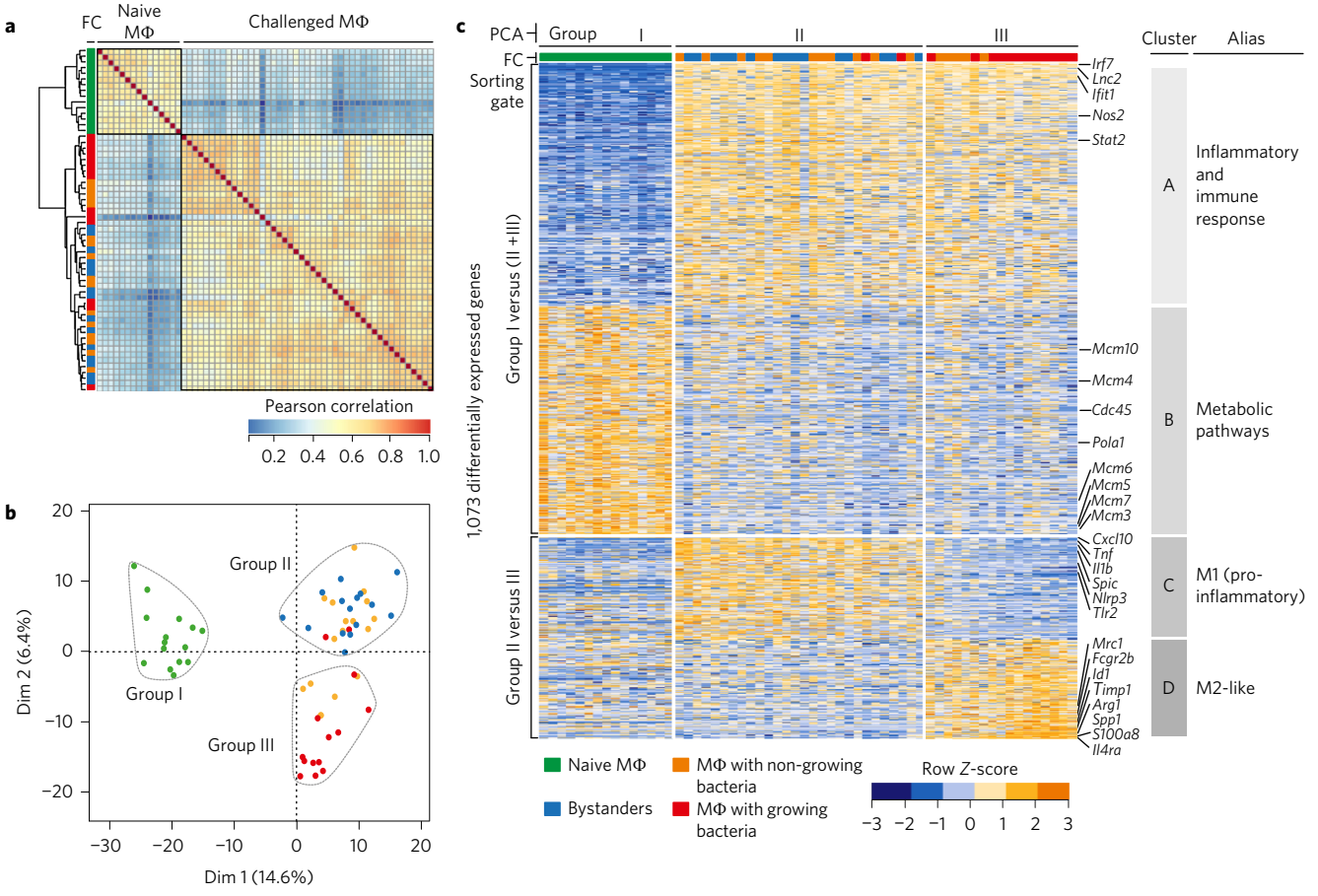

Figure 2 | Single-cell RNA-seq profiling reveals specific transcriptional signatures associated with MΦs containing non-growing or growing *Salmonella*. **a**, Gene expression correlation between individual cells from distinct sorting fractions. The heatmap displays Pearson's *r* values clustered on the *y* axis by the euclidian method (dendrogram). This matrix revealed the existence of two major groups of cells corresponding to naive and challenged cells. Colours indicate cellular identities as inferred from the original FC gates. **b**, PCA allowed for a refinement of the group of challenged cells into groups II and III (group I: naive MΦs). Each dot represents a single cell (60 cells in total). Colours indicate cellular identities as inferred from the original FACS gates. **c**, Transcriptional profiles specific to individual cell groups (60 cells in total). A total of 1,073 genes were differentially expressed, as identified by SCDE ($P < 0.01$; Supplementary Fig. 7) among the three cell groups identified in **b**, and were grouped into individual gene clusters (A–D) and plotted as a heatmap. Clusters A and B are derived by comparing naive cells to challenged cells, labelled respectively cell groups I and II plus III on the PCA map. Clusters C and D are derived by comparing cell groups II and III (Supplementary Fig. 7 and Supplementary Table 2b–d).

forming a distinct cluster from MΦs infected with growing bacteria and a minority of MΦs with non-growing bacteria. The existence of three clusters was further validated by *k*-means clustering embedded in the rare cell type identification (RaceID) algorithm (see Methods and Supplementary Fig. 4b). We also confirmed our clustering using an alternative and complementary approach, PAGODA (see Methods), extracting the major gene sets underlying cellular heterogeneity (Supplementary Fig. 5). Genes related to 'cell cycle' and 'DNA replication' (gene ontology (GO) terms) are highly downregulated in challenged cells (Fig. 2c and Supplementary Fig. 5), suggesting that cell cycle is not required for specific correction of the clustering. Furthermore, an additional PCA on all infected MΦs revealed a similar segregation of the two populations of cells containing growing and non-growing *Salmonella*, indicating that they have dissimilar transcriptome signatures (Supplementary Fig. 6).

To characterize the functional state of the different subpopulations, we determined the set of genes specific to each cell group using the 'single-cell differential expression' (SCDE) approach (see Methods). Briefly, groups I, II and III were compared in a pairwise manner, and all differentially expressed genes ($P < 0.01$) were used for further analysis (Supplementary Table 2a–c and Supplementary Fig. 7) and displayed on a heatmap (Fig. 2c). Overall, 1,073 genes were differentially expressed among the three groups. These genes formed four gene clusters termed A to D (Fig. 2c and Supplementary Table 2d). Clusters A and B (385 and 369 genes, respectively) regroup differentially expressed genes between naive and challenged cells, whereas genes of clusters C and D (161 and 158 genes, respectively) were specifically expressed in group II or III MΦs.

Global host responses to pathogenic stimuli. On the functional level, only challenged MΦs expressed genes known to be activated in immune cells exposed to PAMPs (ref. 10). For example, genes in cluster A include those of master regulators of an LPS-induced immune response (*Iift1*, *Stat2*, *Irf7*), several chemokines (*Ccl2*, *-3*, *-4*, *-5* and *-7*), as well as *Lcn2*, which is linked to innate immunity to bacterial infections (Fig. 2c). The two top GO terms for cluster A were 'Immune response' and 'Inflammatory response' ($P < 1 \times 10^{-10}$, Fisher exact test; Fig. 2c). Conversely, gene cluster B, containing genes downregulated in challenged MΦs as compared to naive cells, was enriched for metabolism-associated GO terms ('DNA metabolism', 'cellular metabolic process', $P < 1 \times 10^{-16}$, Fisher exact test; Fig. 2c). This confirms a previous report showing that professional phagocytes respond to internalized microbes by reorganizing their metabolic activities[15], in addition to mounting immune responses.

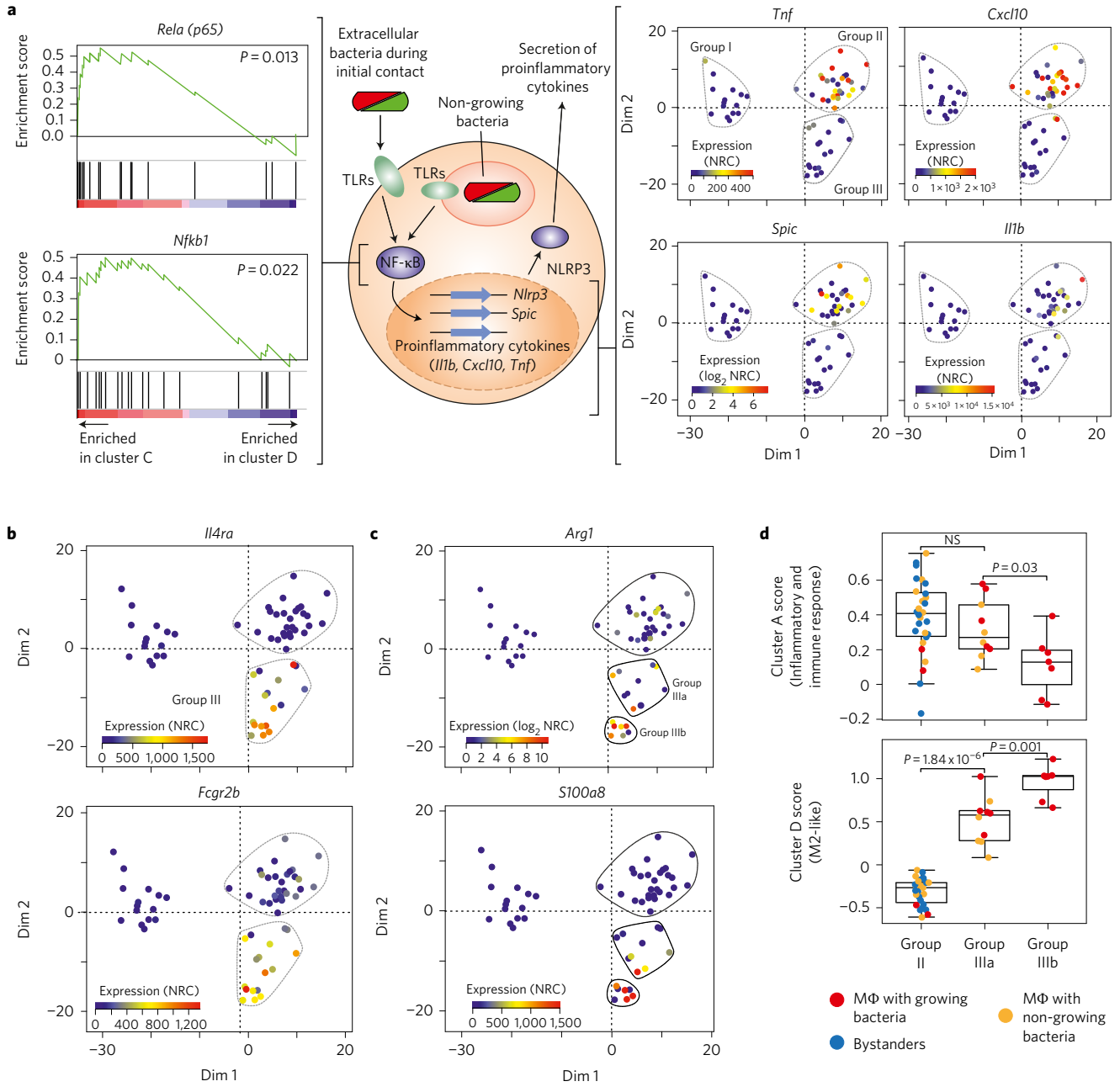

Figure 3 | Different bacterial loads correlate with divergent macrophage polarization transcription programs. a, MΦs containing non-growing bacteria and bystander cells show characteristics of a proinflammatory immune response. Left: GSEA reveals a significant enrichment of *Rela* and *Nfkb1* in cells from group II (that is, MΦs with non-growing bacteria and bystander cells, 28 cells) compared to group III cells (mostly MΦs with growing bacteria, 17 cells). The scheme in the middle outlines the proinflammatory response towards extracellular and intracellular *Salmonella* sensing by the host (adapted from ref. 11). Right: Single-cell expression levels (NRC, normalized read counts) for selected host genes (*Tnf*, *Cxcl10*, *Spic*, *Il1b*) were plotted onto the PCA map shown in Fig. 2b. This representation highlights the elevated expression levels of these genes in cells from group II (MΦs with non-growing bacteria and bystander cells), but also indicates the high extent of variability between individual members of the group. **b**,**c**, MΦs with growing bacteria show a spectrum of M2-like activation. Single-cell gene expression data for *Il4ra*, *Fcgr2b* (**b**) and *Arg1*, *S100a8* (**c**) are plotted on PCA maps. Expression of *Arg1* and *S100a8* (**c**) was specific to a subpopulation within group III, thus allowing for the refinement of these cells into groups IIIa (low *Arg1*, *S100a8*) and IIIb (high *Arg1*, *S100a8*). **d**, Expression scores of genes from clusters A (immune and inflammatory response) and D (M2-like) as identified in Fig. 2c are plotted for the different subgroups of challenged cells. The box plots give the minimum, first quartile, median, third quartile and maximum of the data. An unpaired two-sample *t*-test was applied to obtain *P*. NS, not significant. TLRs, toll-like receptors.

Proinflammatory profile of MΦs with non-growing bacteria matches that of bystander cells. Both the above PCA and SCDE analyses indicated that MΦs with defined bacterial loads exhibit specific mRNA expression programs (Fig. 2b,c). To better understand these mRNA profiles, we performed gene set enrichment analysis (GSEA; see Methods). Comparison of clusters C and D (that is, genes specifically expressed in group II or group III MΦs; Fig. 2c) revealed an over-representation of the prototypical proinflammatory transcription factor NF-κB (*p65* and *nfkb1*)[16] and of several of its target genes in cluster C (*Tnf*, *Cxcl10*, *Il1b*) (Fig. 3a). Additionally, the mRNAs of NLRP3 (NOD-like receptor family, pyrin domain-containing 3; an NF-κB

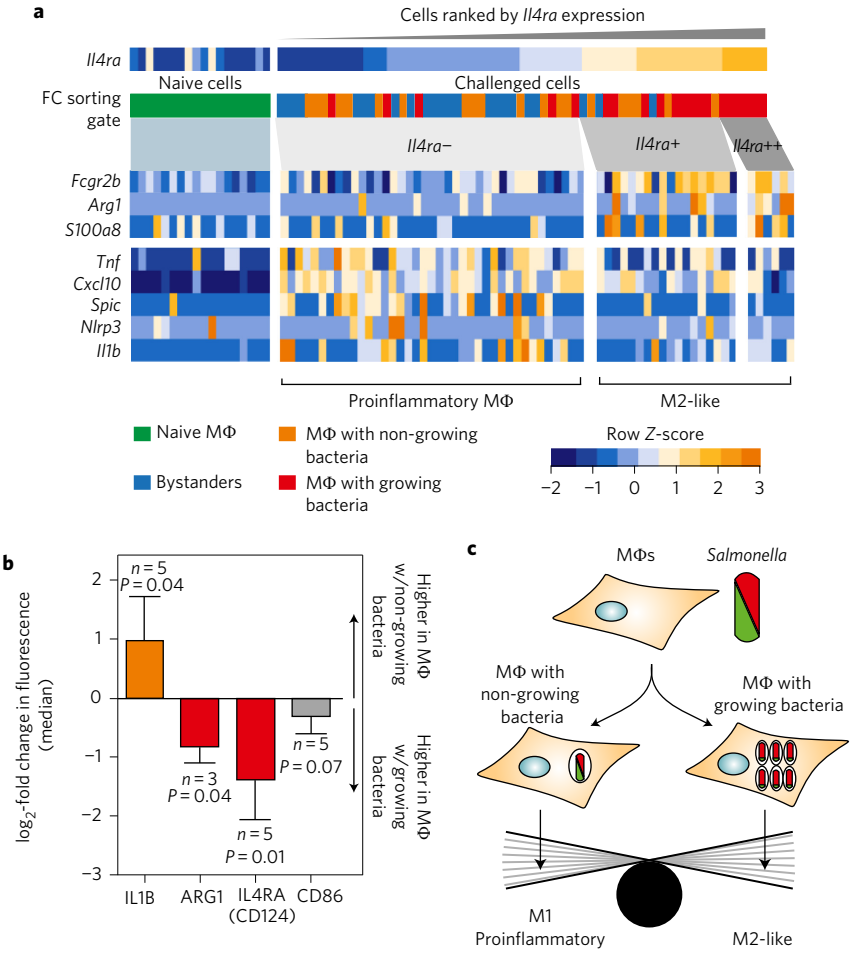

Figure 4 | Bacterial growth triggers a different MΦ polarization pattern. a, 18 naive and 63 challenged cells (23 bystanders, 20 MΦs with non-growing and 20 MΦs with growing *Salmonella*) were isolated and sequenced, and are ranked by increasing *Il4ra* expression (log-transformed, row *Z*-score). The original FC gate of each cell is given at the top (colour code at the bottom). Again, three groups of challenged cells could be distinguished (*Il4ra*−, *Il4ra*+ and *Il4ra*++). The heatmap displays the expression of genes associated with MΦs containing growing and non-growing bacteria. **b**, MΦs were infected with *Salmonella* for 20 h before harvest, antibody staining and FC ($n > 30{,}000$ cells). Cells were gated depending on the growth status of the bacteria as indicated in **a**, and the detected levels of IL1B, ARG1, IL4RA (CD124) and CD86 were compared among the different MΦ populations. Error bars (standard deviation) are calculated for *n* independent biological replicates as indicated. A two-tailed one-sample *t*-test was applied to obtain *P*. **c**, Working model depicting the correlation between different MΦ activation programs and differentially growing *Salmonella*. Two distinct functional populations were found: a proinflammatory activation state dominated in MΦs with non-growing *Salmonella* and bystander cells, whereas an anti-inflammatory (M2-like) state prevailed in MΦs with growing bacteria.

dependent gene) and TLR2 (Toll-like receptor 2; upstream of NF-κB), which are both activated by extracellular and intracellular PAMP sensing and encode proteins of inflammatory signalling cascades (reviewed in ref. 11), were enriched in cluster C (Fig. 3a and Supplementary Fig. 8a). Intriguingly, the activation of all these proinflammatory factors is classically linked to the M1 polarization program of MΦs (ref. 16). These proinflammatory genes were heterogeneously expressed in group II MΦs (Fig. 3a and Supplementary Fig. 8b). We compared cluster A and C to a previous single-cell RNA-seq study[3] of the host response at a much earlier time p.i. (4 h as compared to 20 h in our study) in cells containing only one *Salmonella*. Nonetheless, we identified an overlap in both data sets for almost half of the differentially expressed genes in our clusters A and C (257 of 546 ≈ 47%; Supplementary Table 2e). Moreover, analysis of the overlapping genes revealed a predominance of GO terms related to immune response and proinflammatory response (Supplementary Table 2e).

In addition to proinflammatory genes, some group II MΦs expressed *Spic* encoding the transcription factor Spi-C (Fig. 3a). Spi-C is the key driver of the development of splenic red pulp MΦs, specialized cells in phagocytosing red blood cells[17]. Intriguingly, haemophagocytes are considered a favourable niche for *Salmonella* during chronic host infection[18], because they can acquire a nutrient-rich intracellular milieu by engulfing other cell types. Taken together, our single-cell RNA-seq analysis indicates that the majority of MΦs with restrained bacterial proliferation display a proinflammatory M1 polarization program. The profiles of these MΦs containing non-growing bacteria were indistinguishable from those of bystander cells, suggesting that intracellular non-growing bacteria do not trigger additional immune signalling than extracellular PAMPs; in other words, these bacteria evade recognition by intracellular immune receptors.

MΦs with growing *Salmonella* exhibit M2 expression markers. Next, to understand how MΦs respond to growing bacteria, we inspected gene cluster D, which is highly specific to group III MΦs (Fig. 2c). This cluster was characterized by genes for cell surface Fc gamma receptors (for example, *Fcgr2b* (*CD32*)), IL-4 receptor-α (*Il4ra*), arginase 1 (*Arg1*) and MΦ mannose receptor 1 (*Mrc1*, also known as CD206) (Figs 2c and 3b,c), all of which are hallmark genes for M2-type MΦs (refs 16,19,20).

L-arginine catabolism via Arg1 is critical for the survival of intracellular bacteria, because it counteracts the production of the highly toxic nitric oxides[21]. Similarly, MRC1 plays a role in the sensing of intracellular bacteria[20]. Thus, to test whether *Arg1* and *Mrc1* expressing MΦs (Fig. 2c and Supplementary Fig. 9) indeed constitute a distinct functional subgroup, we computed a 'cluster score' that represents the expression of all genes in a cluster in one number (Fig. 2c; see Methods). This calculation confirmed that cells differing in the expression of prototypical *Arg1* and *Mrc1* also exhibited a reduced activity of the 'inflammatory and immune response' (cluster A) and an increased 'M2-like status' (cluster D) (Fig. 3d). Based on these observations, group III may indeed be further divided into two distinct subpopulations: group IIIa is a mixed MΦ population containing both non-growing and growing *Salmonella*, and group IIIb MΦs contain only growing *Salmonella* (Fig. 3c,d and Supplementary Fig. 9).

We then analysed the expression levels of other immune genes expressed along with M2 markers in cells from group IIIb (that is, in cells with growing *Salmonella*) and observed high expression levels of an antimicrobial calprotectin (*S100a8*)[22] (Fig. 3c) and *Spp1* (*Secreted Phosphoprotein 1*; a.k.a. *Osteopontin* or *Eta-1*) (Supplementary Fig. 9). Both the calprotectin and SPP1 proteins play major roles in host protection against intracellular bacteria[23]. In addition, among the genes specific to this group of cells (gene cluster D) were inhibitor of DNA-binding (ID) genes (*Id1*, *Id2*) (Supplementary Fig. 6). Finally, the chemokine gene *Ccl8* and several poorly characterized genes such as *Timp1*, *CD244*, *Nxpe5*, *Gpr35*, *Ltb4r1* and *CD164* were co-expressed together with M2 marker genes (Supplementary Fig. 9). Our observations indicate that M2 polarization of MΦs happens as a result of infection, as such M2 markers are absent from naive MΦs, and correlates with bacterial proliferation. When we compared the specific host response for MΦs with growing bacteria (cluster D) to the aforementioned previous single-cell RNA-seq results[3], we find only 8% overlap (13 of 158 genes; Supplementary Table 2e). Most importantly, all M2 polarization markers (for example, *Arg1*, *Il4ra* and *Mrc1*) were silent in those early infection data[3], which further supports a view that *Salmonella* growth drives M2 polarization.

Similarities between MΦs infected with differentially growing *Salmonella*. The ten MΦs of group IIIa were composed of cells containing growing *Salmonella* and others harbouring non-growing *Salmonella* (each five cells) (Fig. 3c and Supplementary Fig. 9). Further inspection indicated that this minority of MΦs containing non-growing bacteria escaped a proinflammatory fate and rather exhibited the M2-like transcriptional signature associated with a more hospitable environment for bacteria, such as that seen for cells from group IIIb. However, gene expression analysis showed a global decrease of 'M2-like score' compared to group IIIb MΦs (Fig. 3d), with the previously discussed M2 marker genes *Arg1* and *Mrc1* being heterogeneously expressed (Fig. 3c and Supplementary Fig. 9). We interpret these cells to represent intermediates of the two above-described M1 and M2 states.

Independent validation of the transcriptomic differences. We sought to independently validate the differential expression of marker genes of the intracellular proliferation phenotype of *Salmonella*. To this end, a second replicate experiment of infection, cell sorting and single-cell RNA-seq was performed. After quality filtering (as above), this experiment comprised a total of 81 cells: 18 naive MΦs, 23 bystanders, 20 MΦs containing non-growing bacteria, and 20 MΦs containing growing bacteria (Fig. 4a and Supplementary Fig. 10). Unsupervised PCA confirmed the initial pattern such that challenged cells segregated into two groups: one group comprising mostly bystanders and MΦs containing non-growing bacteria, and the other composed predominantly of MΦs with growing bacteria (Supplementary Fig. 10).

We classified the cells according to their level of *Il4ra* expression and confirmed that *Il4ra*, identified above to be highly expressed in MΦs infected with growing bacteria (Fig. 3b), can be used as a discrimination marker for the groups of proinflammatory and M2-like MΦs (Fig. 4a). Cells that exhibited the highest levels of *Il4ra* mRNA also showed a high expression of M2 marker genes (*Fcgr*, *Arg1* and *S100a8*). Conversely, proinflammatory genes (*Tnf*, *Cxcl10*, *Il1b* and *Spic*) were again predominantly expressed in MΦs containing non-growing bacteria and bystander cells, thereby confirming our former observations.

Seeking to validate our single-cell RNA-seq results on the protein level, we used a panel of antibodies targeting IL1B and CD86 as M1 markers and IL4RA (CD124) and ARG1 as M2 markers. Flow cytometry (FC) analysis confirmed that, in agreement with the transcriptomics data, the IL1B, IL4RA and ARG1 proteins were differentially expressed in MΦs with non-growing or growing bacteria, whereas CD86 remained unchanged (Fig. 4b). Moreover, tracking levels of IL4RA over time showed that this M2 marker accumulates as bacterial growth occurs (Supplementary Fig. 11). In addition, this apparent polarization of MΦs was dependent on bacterial growth rather than bacterial numbers: MΦs containing multiple non-growing bacteria did not display the polarization of MΦs containing growing bacteria (Supplementary Fig. 12).

Discussion

Our results show that individual MΦs carrying growing or non-growing *Salmonella* have divergent gene activation programs, as summarized in Fig. 4c. Most MΦs with non-growing bacteria resemble M1 polarization and are very similar to bystander cells. This suggests that non-growing bacteria might be in a stealth mode, in which they do not trigger additional immune recognition by intracellular receptors. We do caution, however, that a potentially saturating activation by LPS sensing may mask a differential transcriptional response for bystanders and non-growing bacteria at early time points. More importantly, we observe that high bacterial loads correlate with an M2-like anti-inflammatory expression program. Generally, such a correlation with M2 polarization is unexpected, because the common response of MΦs to bacterial and viral infections is the upregulation of genes involved in M1 polarization[24]. However, our observation is consistent with previous reports showing preferential replication of *Salmonella* in certain types of isolated MΦs, that is, haemophagocytic[18] or IL-4-stimulated MΦs[12], as well as *Salmonella*'s presence in anti-inflammatory MΦs in the spleen and mesenteric lymph nodes in a murine model of chronic infection[12,25]. This notwithstanding, our growth-rate-correlated results pose the exciting question of whether the observed heterogeneity in MΦ polarization is the cause or consequence of fast and slow *Salmonella* growth. Given that we see little if any M1/M2 polarization in naive MΦs (Figs 2b and 4a), it is tempting to speculate that the intracellular *Salmonella* drive those MΦs unable to clear the infection away from the hostile M1 to the more permissive M2 polarization state. Such reprogramming of infected MΦs would constitute a more dramatic manipulation than a previously known interference of *Salmonella* with M1 polarization through bacterial inhibition of the relocalization of NADPH oxidase to the MΦ phagosome[26].

Of the other host states, we observe that some M1-activated MΦs with non-growing bacteria (and a few cells with growing bacteria) seem to differentiate into red pulp MΦs. Because haemophagocytes constitute a rich source of nutrients supporting chronic infection[18], such an acquisition of novel phenotypic properties by proinflammatory MΦs might also support the persistence of non-growing bacteria. In addition, *Salmonella* has been shown to reside in red pulp MΦs outside any visible lesion, making the pathogen 'invisible' to immune surveillance[27].

Altogether, our present results support the emerging idea that bacteria use host genome plasticity to subvert infected cells. *Salmonella* was recently shown to transform follicle-associated intestinal epithelial cells into M cells[28]. Similarly, *Mycobacterium leprae* reprograms non-immune Schwann cells into progenitor/stem-like cells for them to serve as a vehicle for bacterial dissemination[29]. However, although many bacteria use MΦs for replication, links to MΦ polarization have so far been indirect, for lack of a suitable model to track bacterial growth[2]. The *Salmonella* model used here promises to reveal bacterial molecules and host mechanisms whereby intracellular bacteria modulate MΦ polarization for their own benefit. Although current single-cell RNA-seq methods are restricted to profiling eukaryotic transcripts[9], technical improvement of the so-called Dual RNA-seq technique, which profiles gene expression in host and bacteria simultaneously[30], may soon permit the correlation of the discrete activities of bacterial virulence factors with host cell heterogeneity and reprogramming.

Methods

***In vitro* infection protocol and single-cell sorting.** Bone marrow-derived MΦs were obtained by isolation and differentiation of the marrow of the hindleg bones of 8- to 12-week-old C57/BL6 wild-type male mice (Janvier Laboratories). Briefly, a bone marrow suspension in X-Vivo 15 medium (Lonza), supplemented with 10% fetal bovine serum (FBS superior, Biochrom) and 10% L929 conditioned DMEM medium (herein called 'differentiation medium') was seeded into six-well plates at a density of 10^6 leukocyte cells in 2 ml per well. Cells were incubated at 37 °C in a CO_2 incubator with a humidified atmosphere for up to 7 days. At day 3, the cultures were supplemented with 0.3 volumes of fresh differentiation medium. At day 7, cells were collected and seeded into six-well tissue culture plates at a density of 1×10^6 cells per well in a fresh 2 ml differentiation medium.

S. Typhimurium SL1344 strain harbouring the pFCcGi plasmid[31] was used for FD (ref. 6). The strain was cultured for 16 h at 37 °C in a minimum medium MgMES supplemented with 0.2% arabinose and 50 μg ml^{-1} ampicillin. Before infection, 180 μl of bacteria culture were opsonized for 20 min at room temperature by supplementing 20 μl of mouse serum and 80 μl of differentiation medium.

MΦs (see above) were infected with *Salmonella* at a multiplicity of infection of 50. Plates were centrifuged for 5 min at 110*g* to encourage contact between bacteria and MΦs. After 30 min, cells were washed twice with room-temperature PBS. The medium was replaced with the differentiation medium (see above) supplemented with 100 μg ml^{-1} gentamicin for 1 h to eliminate extracellular bacteria. Gentamicin concentration was reduced to 20 μg ml^{-1} for the remainder of the infection. At 20 h post-uptake, cells were resuspended in 1 ml of ice-cold PBS by scraping from each well and single-cell suspension was prepared by passing cells through a 30 μm pre-separation filter (Miltenyi Biotec).

Single cells were immediately sorted using a FACSaria III (BD Biosciences; device: 96-well plate; precision: single-cell; nozzle: 100 μm). Forward-scatter area (FCS-A) versus side-scatter area (SSC-A) was used to gate out damaged cells, dead cells were stained with 1 μl of saturated propidium iodide (PI), and dead PI+ cells were gated out.

Single cells were sorted into different wells from a 48-well plate (Brand) filled with 3 μl of a hypotonic lysis buffer (2.7 μl of 0.2% Triton X-100 (Sigma) in RNAse-free water (Life Technologies) and 0.3 μl of SUPERase In (Life Technologies)). Single sorted cells were immediately chilled in a −20 °C prepared isofreeze rack and processed for single-cell RNAseq or stored at −80 °C.

Single-cell RNAseq and tagmentation protocol. Single-cell libraries were prepared as reported in the Smart-seq2 protocol[14,32] with some technical adaptation: 1 μl of ERCC spike-Mix 1 (Life Technologies) diluted at $1:10^6$ in RNAse-free water was added after cell sorting to the lysis buffer; template-switching oligo (TSO) was modified to add isomeric nucleotide at the 5′end to minimize background cDNA synthesis as previously described[33] (5′-(iso-dC)(iso-dG)(iso-dC) AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3′, Eurogentec); after first strand synthesis primer–dimers were removed using 11 μl of Ampure XP beads (Beckman Coulter). After cDNA amplification, libraries were quantified using Qubit Hs Assay (Life Technologies) and quality of libraries was checked using Bioanalyser (Agilent). Using the modified protocol described in ref. 34, 1 ng of cDNA was subjected to a tagmentation-based protocol (Nextera XT, Illumina) using one-quarter of the recommended volumes, 10 min for tagmentation at 55 °C and 1 min extension time during PCR. After PCR and cDNA purification, cDNA was purified and resuspended in 15 μl of elution buffer (EB) (Qiagen). Finally, libraries were pooled (16 libraries for MiSeq 96 libraries for NextSeq) and sequencing was performed in paired-end mode for 2×75 cycles using Illumina's MiSeq or NextSeq.

Single-cell RNA-seq read mapping and expression quantification. RNA-seq data are deposited in Gene Expression Omnibus GSE79363. After demultiplexing, an initial quality assessment was performed using FastQC (v 0.11.3). Illumina adapters and Smart-seq2 PCR primer (5′-AAGCAGTGGTATCAACGCAGAGT-3′) were removed using cutadapt (v1.3)[35]. Reads below a phred score of 20 were trimmed using the program fastq_quality_trimmer from the FASTX toolkit (v 0.0.13). After quality trimming and adapter clipping, unmatched paired end reads were removed using in-house scripts.

Trimmed reads were mapped to the mouse genome reference sequence (GRCm38.p3) and 92 ERCC spike-in sequences. Alignment was performed using Bowtie2 (v 2.2.4)[36] and Tophat (v 2.0.13)[37] with default settings. The incorrectly annotated transcript ENSMUSG0000092329 was removed from the mouse annotation file. Read counts of each gene were determined using the HTSeq program[38]. FPKM values were calculated using Cufflinks (v 2.2.1)[39] using the no-effective-length-correction and compatible-hits-norm options.

Quality control and cell normalization. Single cells with total number of trimmed reads lower than 400,000, an *Actb* gene expression lower than 1,000 (in FPKM) and/or percentage of reads dedicated to ERCC larger than 70% were discarded. We then applied a filtering step to remove genes not expressed or expressed in fewer than two cells. Finally, a size factor for each cell was estimated based on technical variations of the ERCC spike-ins and later used to normalize all the endogenous genes[40].

Variable gene analysis and PCA. To infer the genes that were variable above the technical noise, a linear regression model was applied to fit the relationship between the squared coefficient of variation (CV^2) with the mean expression of ERCC spike-ins as previously introduced[41]. All single cells that passed the filter were subjected to a PCA (FactomineR R package[42]) using the variable genes. The first four principle component scores of each cell in each gene cluster were retrieved with the 'indcoord' option.

Single-cell differential expression and pathway overdispersion analysis. To identify the differential expression signature between groups of cells identified by PCA, a Bayesian method for SCDE was used[43]. A *Z* score calculated by SCDE was further transformed to an empirical *P* value (cutoff $P < 0.01$) to determine significant genes (Supplementary Fig. 7). The GOs with coordinated variability among subpopulations were detected by the PAGODA algorithm[44], which was implemented in the SCDE package (Supplementary Fig. 5).

GO analysis. The GO of a specific gene list was analysed using the topGO package in R Bioconductor. Briefly the GO terms of each gene were retrieved using the annFun function, and then the ontology of each gene was classified into three categories: molecular function, biological process and cellular component. Enrichment analysis was performed to evaluate the significance of the terms compared to all the mouse genes using the Fisher exact test. The calculated *P* values were further corrected using the Benjamini–Hochberg method. The subgraph of significant GO terms was visualized using the showSigofNodes function.

Correlation plots, heatmaps and cluster score. Cell–cell Pearson correlation were performed with the 'cor' function in R, and further plotted using the pheatmap R package with clustering methods set to 'average'. The normalized expression of each gene was transformed into a log value (log(expression + 1)), and row scaled before generating a heatmap using the heatmap.2 function in the gplots R package. To summarize the expression of each gene cluster into a numeric single value for each single cell, a 'cluster score' was calculated as the mean expression for every gene present in the cluster. The cluster score was visualized using the boxplot and stripchart R function, and the jitter method was used to separate cells.

Gene set enrichment analysis. Mouse-curated pathway and transcription factors were associated with a set of gene lists similar to a gene cluster. GseaPreranked function in program Gene set enrichment analysis (GSEA)[45] was used to calculate an enrichment score for each gene set sorted based on log fold change with the following parameters ('-collapse false -mode Max_probe -norm meandiv -nperm 1000 -scoring_scheme weighted -include_only_symbols true -make_sets true -rnd_seed timestamp -set_max500 -set_min 15'). The curated pathways and transcription factor target genes annotation were retrieved from GSKB in the Bioconductor experiment package (v 3.2).

RaceID. The number of major clusters was inferred by gap statistic and further a *k*-means clustering was applied to the similarity matrix of these clusters using RaceID[46]. The outlier cells within each cluster was identified by the method 'findoutliers' with default values.

Macrophage polarization detection. Bone marrow-derived MΦs were recovered at different times before or after infection (as described above) and fixed in 500 μl of 3% paraformaldehyde for 15 min. Cells were then washed once in 500 μl cold PBS. For IL1B and ARG1 labelling, cells were permeabilized in 0.1% saponin for 30 min. Labelling was carried out in 150 μl PBS supplemented with 10% horse serum (and 0.1% saponin when appropriate) and the conjugated antibodies (BD Biosciences) for 45 min at room temperature. Cells were washed once in PBS before being analysed

on a Fortessa Flow cytometer (BD Biosciences). The antibody panel consisted of IL4RA (CD124)-BV421, CD86-PE/Cy7, IL1B-PE and ARG1-APC.

Data availability. RNA-seq data supporting the findings of this study have been deposited at Gene Expression Omnibus under accession no. GSE79363.

Received 19 March 2016; accepted 16 September 2016; published 14 November 2016

References

1. Bumann, D. Heterogeneous host–pathogen encounters: act locally, think globally. *Cell Host Microbe* **17,** 13–19 (2015).
2. Price, J. V. & Vance, R. E. The macrophage paradox. *Immunity* **41,** 685–693 (2014).
3. Avraham, R. *et al.* Pathogen cell-to-cell variability drives heterogeneity in host immune responses. *Cell* **162,** 1309–1321 (2015).
4. Monack, D. M., Mueller, A. & Falkow, S. Persistent bacterial infections: the interface of the pathogen and the host immune system. *Nat. Rev. Microbiol.* **2,** 747–765 (2004).
5. Claudi, B. *et al.* Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy. *Cell* **158,** 722–733 (2014).
6. Helaine, S. *et al.* Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. *Science* **343,** 204–208 (2014).
7. Helaine, S. *et al.* Dynamics of intracellular bacterial replication at the single cell level. *Proc. Natl Acad. Sci. USA* **107,** 3746–3751 (2010).
8. Jenner, R. G. & Young, R. A. Insights into host responses against pathogens from transcriptional profiling. *Nat. Rev. Microbiol.* **3,** 281–294 (2005).
9. Saliba, A. E., Westermann, A. J., Gorski, S. A. & Vogel, J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* **42,** 8845–8860 (2014).
10. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510,** 363–369 (2014).
11. Keestra-Gounder, A. M., Tsolis, R. M. & Bäumler, A. J. Now you see me, now you don't: the interaction of *Salmonella* with innate immune receptors. *Nat. Rev. Microbiol.* **13,** 206–216 (2015).
12. Eisele, N. A. *et al. Salmonella* require the fatty acid regulator PPARδ for the establishment of a metabolic environment essential for long-term persistence. *Cell Host Microbe* **14,** 171–182 (2013).
13. Kreibich, S. & Hardt, W. D. Experimental approaches to phenotypic diversity in infection. *Curr. Opin. Microbiol.* **27,** 25–36 (2015).
14. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10,** 1096–1098 (2013).
15. Eisenreich, W., Heesemann, J., Rudel, T. & Goebel, W. Metabolic host responses to infection by intracellular bacterial pathogens. *Front. Cell Infect. Microbiol.* **3,** 24 (2013).
16. Lawrence, T. & Natoli, G. Transcriptional regulation of macrophage polarization: enabling diversity with identity. *Nat. Rev. Immunol.* **11,** 750–761 (2011).
17. Kohyama, M. *et al.* Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature* **457,** 318–321 (2009).
18. Nix, R. N., Altschuler, S. E., Henson, P. M. & Detweiler, C. S. Hemophagocytic macrophages harbor *Salmonella enterica* during persistent infection. *PLoS Pathog.* **3,** e193 (2007).
19. Guilliams, M., Bruhns, P., Saeys, Y., Hammad, H. & Lambrecht, B. N. The function of Fcγ receptors in dendritic cells and macrophages. *Nat. Rev. Immunol.* **14,** 94–108 (2014).
20. Martinez, F. O., Helming, L. & Gordon, S. Alternative activation of macrophages: an immunologic functional perspective. *Annu. Rev. Immunol.* **27,** 451–483 (2009).
21. Bronte, V. & Zanovello, P. Regulation of immune responses by L-arginine metabolism. *Nat. Rev. Immunol.* **5,** 641–654 (2005).
22. De Jong, H. K. *et al.* Expression and function of S100A8/A9 (calprotectin) in human typhoid fever and the murine *Salmonella* model. *PLOS Negl. Trop. Dis.* **9,** e0003663 (2015).
23. Ashkar, S. *et al.* Eta-1 (osteopontin): an early component of type-1 (cell-mediated) immunity. *Science* **287,** 860–864 (2000).
24. Benoit, M., Desnues, B. & Mege, J. L. Macrophage polarization in bacterial infections. *J. Immunol.* **181,** 3733–3739 (2008).
25. McCoy, M. W., Moreland, S. M. & Detweiler, C. S. Hemophagocytic macrophages in murine typhoid fever have an anti-inflammatory phenotype. *Infect. Immun.* **80,** 3642–3649 (2012).
26. Vazquez-Torres, A. *et al.* Salmonella pathogenicity island 2-dependent evasion of the phagocyte NADPH oxidase. *Science* **287,** 1655–1658 (2000).
27. Burton, N. A. *et al.* Disparate impact of oxidative host defenses determines the fate of *Salmonella* during systemic infection in mice. *Cell Host Microbe* **15,** 72–83 (2014).
28. Tahoun, A. *et al. Salmonella* transforms follicle-associated epithelial cells into M cells to promote intestinal invasion. *Cell Host Microbe* **12,** 645–656 (2012).
29. Masaki, T. *et al.* Reprogramming adult Schwann cells to stem cell-like cells by leprosy bacilli promotes dissemination of infection. *Cell* **152,** 51–67 (2013).
30. Westermann, A. J. *et al.* Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions. *Nature* **529,** 496–501 (2016).
31. Figueira, R., Watson, K. G., Holden, D. W. & Helaine, S. Identification of *Salmonella* pathogenicity Island-2 type III secretion system effectors involved in intramacrophage replication of *S. enterica* serovar typhimurium: implications for rational vaccine design. *mBio* **4,** e00065-13 (2013).
32. Picelli, S. *et al.* Full-length RNA-seq from single cells using smart-seq2. *Nat. Protoc.* **9,** 171–181 (2014).
33. Kapteyn, J., He, R., McDowell, E. T. & Gang, D. R. Incorporation of non-natural nucleotides into template-switching oligonucleotides reduces background and improves cDNA synthesis from very small RNA samples. *BMC Genomics* **11,** 413 (2010).
34. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344,** 1396–1401 (2014).
35. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).
36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
37. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25,** 1105–1111 (2009).
38. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
39. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with topHat and cufflinks. *Nat. Protoc.* **7,** 562–578 (2012).
40. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
41. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10,** 1093–1095 (2013).
42. Lê, S., Josse, J. & Husson, F. Factominer: an R package for multivariate analysis. *J. Stat. Soft.* **25,** 1–18 (2008).
43. Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11,** 740–742 (2014).
44. Fan, J. *et al.* Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat. Methods* **13,** 241–244 (2016).
45. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
46. Grün, D. *et al.* Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525,** 251–255 (2015).

Acknowledgements

The authors thank T. Achmedov, V. McParland, H. Merkert and B. Plaschke for technical support. A.-E.S. was supported by the PostDoc Plus program of the University of Würzburg. A.J.W. was the recipient of an Elite Advancement PhD stipend from Universität Bayern e.V. S.H. was supported by an MRC Career Development Award (MR/M009629/1). D.A.C.S. was the recipient of an EMBO postdoctoral fellowship (ALTF 441-2015).

Author contributions

A.-E.S., A.J.W. and J.V. designed and A.-E.S. performed the experiments. L.L. and S.A. performed bioinformatics analysis. S.H. and D.A.C.S. performed flow cytometry experiments. A.J.W., S.H. and L.N.S. provided reagents. A.-E.S., A.J.W., S.H. and J.V. wrote the manuscript.

Additional information

Supplementary information is available for this paper. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.V.

Competing interests

The authors declare no competing financial interests.